

Department of Economics

Working Paper Series

## Bootstrap Unit Root Tests in Models with GARCH(1,1) Errors

09-001 | Nikolay Gospodinov  
Concordia University  
Ye Tao  
Concordia University



UNIVERSITÉ

Concordia

UNIVERSITY

# BOOTSTRAP UNIT ROOT TESTS IN MODELS WITH GARCH(1,1) ERRORS\*

Nikolay Gospodinov<sup>†</sup>

Concordia University and CIREQ

Ye Tao<sup>‡</sup>

Concordia University

January 2009

Revised: May 2009

## Abstract

This paper proposes a bootstrap unit root test in models with GARCH(1,1) errors and establishes its asymptotic validity under mild moment and distributional restrictions. While the proposed bootstrap test for a unit root shares the power enhancing properties of its asymptotic counterpart (Ling and Li, 2003), it offers a number of important advantages. In particular, the bootstrap procedure does not require explicit estimation of nuisance parameters that enter the distribution of the test statistic and corrects the substantial size distortions of the asymptotic test that occur for strongly heteroskedastic processes. The simulation results demonstrate the excellent finite-sample properties of the bootstrap unit root test for a wide range of GARCH specifications.

*JEL Classification:* C12, C15, C22.

*Keywords:* Unit root test; GARCH; Bootstrap.

---

\*We would like to thank the Editor and two anonymous referees for helpful comments and suggestions. The first author gratefully acknowledges financial support from FQRSC, IFM2 and SSHRC.

<sup>†</sup>Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada; tel: (514) 848-2424 (ext.3935), fax: (514) 848-4536, email: [nikolay.gospodinov@concordia.ca](mailto:nikolay.gospodinov@concordia.ca).

<sup>‡</sup>Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8 Canada; fax: (514) 848-4536, email: [tao627@hotmail.com](mailto:tao627@hotmail.com).

# 1 Introduction

The simultaneous presence of high persistence and conditional heteroskedasticity is a common characteristic of many economic time series. The stark differences between the long-run behavior of nonstationary and stationary processes and their implications for economic analysis led to the development of a large class of unit root tests with good size and power properties. While the limiting theory for possibly unit root processes has been established under fairly general conditions, including some types of time-varying volatility, the explicit modeling of the higher-order dynamics is often expected to improve the efficiency of the conditional mean estimates and the power of the tests. For instance, a strand of literature that emerged recently (Ling and Li, 1998, 2003; Ling, Li and McAleer, 2003; Seo, 1999) derives the asymptotic distributions of unit root tests with GARCH errors and demonstrates the power gains of incorporating the GARCH structure into the testing procedure. The form of the asymptotic distribution of the unit root test in this case is a mixture of a Dickey-Fuller (DF) and a standard normal distribution with a mixing coefficient that depends on the degree of conditional heteroskedasticity. As the degree of conditional heteroskedasticity increases (i.e., the sum of the GARCH coefficients approaches one), the standard normal distribution carries more weight and the corresponding smaller critical values give rise to a more powerful testing procedure. Note that the Dickey-Fuller distribution is still valid in the presence of GARCH errors but it is conservative and provides an upper bound for the critical values.

Despite the non-trivial power gains of the unit root tests with GARCH errors (see, for example, Seo, 1999), the applied work with these tests has been very limited. There are several possible reasons why empirical researchers may find these tests not to be particularly appealing. First, they require nonlinear (maximum likelihood) estimation as opposed to OLS estimation for the Dickey-Fuller tests. More importantly, the asymptotic distribution depends on nuisance parameters which involves additional computation for obtaining critical values. Finally, as we show later in the paper, the tests based on asymptotic critical values suffer from substantial size distortions especially for some GARCH parameter configurations that are typically documented in empirical studies with financial time series data.

In this paper we propose a bootstrap method for approximating the finite-sample distributions of unit root tests with GARCH(1,1) errors and establish its asymptotic validity. We extend the

results of Basawa *et al.* (1989, 1991), Ferretti and Romo (1996), Heimann and Kreiss (1996) and Park (2003), among others, to unit root models with conditional heteroskedasticity estimated by maximum likelihood (ML). The implementation of the proposed bootstrap procedure is straightforward and is valid under some fairly weak conditions. In particular, we follow Ling and Li (2003) and derive the consistency of the bootstrap distribution assuming finite second moments and symmetry of the errors. This allows for highly persistent GARCH specifications (with sum of the GARCH parameters arbitrarily close to one) that are commonly estimated in the empirical finance literature. Some related bootstrap results are derived in Gospodinov (2008) in the context of testing for nonlinearity in models with a unit root and GARCH errors.

The finite-sample results demonstrate the excellent size and power properties of the proposed bootstrap test. While the tests based on asymptotic critical values tend to overreject (in some situations, up to 40-50% at 5% significance level), the bootstrap test is always very close to its nominal size regardless of the degree of conditional heteroskedasticity. Furthermore, the power of the bootstrap test that incorporates the GARCH structure of the model exceeds the size-adjusted power of the standard DF test by a substantial margin when the conditional heteroskedasticity is strong.

The properties of the proposed bootstrap test prove to be of great practical importance for identifying the mean reverting behavior in processes with GARCH structure. In our empirical analysis of several U.S. interest rate processes, we show that the DF test does not provide any evidence against the null of a unit root which has important implications about the long-run properties of the data. In contrast, the bootstrap DF-GARCH test tends to reject convincingly the unit root hypothesis due to its superior power properties. This lends support to the mean reverting specification as an underlying process for interest rate dynamics in many economic and finance models.

The rest of the paper is organized as follows. The main model and notation are introduced in Section 2. Section 3 describes the proposed bootstrap procedure and derives its asymptotic validity. Section 4 presents a Monte Carlo simulation experiment that assesses the finite-sample performance of the asymptotic and bootstrap tests and illustrates their usefulness for term structure of interest rates. Section 5 concludes. The proofs of all results in the paper are relegated to the Appendix.

## 2 Model and Notation

Consider the first-order AR process with GARCH(1,1) errors

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} \eta_t \\ h_t &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}, \end{aligned} \tag{1}$$

where  $\phi = 1$  and  $\eta_t \sim iid(0, 1)$ . This model can be generalized to higher-order AR and GARCH processes. For simplicity, we present the results for the first-order model (1) but the limiting representations and the bootstrap procedure can be extended in a straightforward manner (but with more cumbersome notation) to higher-order processes. We can also allow for non-zero mean and linear trend in  $y_t$  in which case the analysis is performed on the demeaned or detrended data. We describe below how the asymptotic and bootstrap procedures need to be modified in the presence of deterministic components.

Let  $\delta = (\omega, \alpha, \beta)$  denote the vector of the unknown GARCH parameters. The parameter of interest is  $\phi$  and the Gaussian quasi-likelihood function of this model is given by

$$L_T(\phi, \delta) = \frac{1}{T} \sum_{t=1}^T l_t(\phi, \delta), \tag{2}$$

where  $l_t(\phi, \delta) = -\frac{1}{2} \ln h_t - \frac{1}{2} \frac{\varepsilon_t^2}{h_t}$ . We follow Ling and Li (2003) and assume that the following conditions are satisfied.

ASSUMPTION 1 *Assume that*

- (a)  $\eta_t \sim iid(0, 1)$ ,  $E(\eta_t^3) = 0$ ,  $E(\eta_t^4) = \kappa < \infty$  for all  $t$ ;
- (b)  $\Psi = \{(\omega, \alpha, \beta) : 0 < \omega_l \leq \omega \leq \omega_u, 0 < \alpha_l \leq \alpha \leq \alpha_u, 0 < \beta_l \leq \beta \leq \beta_u, \alpha + \beta < 1\}$ ;
- (c)  $y_0 = 0$  and  $h_0$  is initialized from its invariant measure.

Assumption 1 imposes some very weak moment and distributional conditions on the error term. The standardized errors are assumed to be symmetric *iid* random variables with a finite fourth moment. The assumed symmetric distribution of  $\eta_t$  may appear restrictive but this allows us to

weaken the moment requirements on the error term  $\varepsilon_t$  (see Ling and Li, 2003). In particular, the limiting results and the validity of the bootstrap procedure are derived assuming the existence of finite second moment of  $\varepsilon_t$  which is satisfied under fairly general conditions on the GARCH parameters. More specifically, the conditions in part (b) ensure that  $E(\varepsilon_t^2) < \infty$  and the processes  $\{h_t\}$  and  $\{\varepsilon_t\}$  are strictly stationary, ergodic and  $\beta$ -mixing with exponential decay (Carrasco and Chen, 2002; Francq and Zakořan, 2006) and allow for strong conditional heteroskedasticity that is typically present in financial data. Part (c) specifies the initialization of the conditional mean and variance functions. Assuming  $y_0$  to be fixed at a different value than zero or to be  $o_p(T^{1/2})$  does not affect the limiting results derived below. Similarly, the asymptotic distributions are invariant to the assumption on the initial condition of  $h$  (Lee and Hansen, 1994; Ling and Li, 2003).

By the block diagonality of the Hessian matrix (Bollerslev, 1986; Ling, Li and McAleer, 2003), the conditional mean and variance parameters can be estimated separately without any efficiency loss. Let  $\hat{\phi}_{LS} = (\sum_{t=1}^T y_{t-1}^2)^{-1} (\sum_{t=1}^T y_t y_{t-1})$  denote the OLS estimator of  $\phi$  and note that  $T(\hat{\phi}_{LS} - 1) = O_p(1)$  under Assumption 1. The parameter vector  $\delta$  can be estimated from the OLS residuals  $\hat{\varepsilon}_t = y_t - \hat{\phi}_{LS} y_{t-1}$  and the corresponding estimates  $\hat{\delta}$  are asymptotically equivalent to the estimates obtained from the true  $\varepsilon_t$ . Then, for some preliminary  $T$ -consistent estimator  $\tilde{\phi}$ , the one-step QMLE estimator of  $\phi$  is given by

$$\hat{\phi}_{ML} = \tilde{\phi} - \left[ \sum_{t=1}^T \frac{\partial^2 l_t(\phi, \hat{\delta})}{\partial \phi^2} \right]_{\phi=\tilde{\phi}}^{-1} \left[ \sum_{t=1}^T \frac{\partial l_t(\phi, \hat{\delta})}{\partial \phi} \right]_{\phi=\tilde{\phi}}$$

and (Ling and Li, 2003)

$$T(\hat{\phi}_{ML} - 1) = - \left[ \frac{1}{T^2} \sum_{t=1}^T \frac{\partial^2 l_t(\phi, \delta)}{\partial \phi^2} \right]_{\phi=1}^{-1} \left[ \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t(\phi, \delta)}{\partial \phi} \right]_{\phi=1} + o_p(1).$$

The OLS estimator  $\hat{\phi}_{LS}$  can be used as an initial preliminary estimator. Then, the iterative estimator that updates the estimates of  $\hat{\delta}$  and  $\hat{\phi}_{ML}$  until convergence is asymptotically equivalent to the full MLE.

Let  $t_{\phi_{LS}=1} = (\sum_{t=1}^T y_{t-1}^2)^{1/2} (T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2)^{-1/2} (\hat{\phi}_{LS} - 1)$  and  $t_{\phi_{ML}=1} = \left[ - \sum_{t=1}^T \frac{\partial^2 l_t(\phi, \delta)}{\partial \phi^2} \right]_{\phi=\hat{\phi}_{ML}, \delta=\hat{\delta}}^{1/2} (\hat{\phi}_{ML} - 1)$  be the  $t$ -statistics of  $H_0 : \phi = 1$  for the OLS and ML estimators, respectively. Let also  $\Rightarrow$  signify weak convergence,  $D[0, 1]$  denote the space of real valued functions defined on the interval  $[0, 1]$  that are right-continuous at each point in  $[0, 1]$  and have finite left limits, and  $B_1(r)$  be a

standard Brownian motion on  $D[0, 1]$ . The following lemma is a restatement of some results in Ling and Li (2003) and Seo (1999).

LEMMA 1 *Suppose that  $\phi = 1$  and Assumption 1 holds. Then, as  $T \rightarrow \infty$*

$$t_{\phi_{LS}=1} \Rightarrow \frac{\int_0^1 B_1(r) dB_1(r)}{\left(\int_0^1 B_1^2(r) dr\right)^{1/2}}, \quad (3)$$

$$t_{\phi_{ML}=1} \Rightarrow \sqrt{\frac{K}{F}} \left[ \rho \frac{\int_0^1 B_1(r) dB_1(r)}{\left(\int_0^1 B_1^2(r) dr\right)^{1/2}} + \sqrt{1 - \rho^2} z \right], \quad (4)$$

where  $\rho = 1/\sigma\sqrt{K}$ ,  $E(h_t) = \sigma^2$ ,  $K = E(1/h_t) + (\kappa - 1)\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2/h_t^2)$ ,  $F = E(1/h_t) + 2\alpha^2 \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2/h_t^2)$  and  $z$  is a standard normal random variable distributed independently of  $B_1(r)$ .

PROOF See Ling and Li (2003) and Seo (1999).

Several interesting observations emerge from the limiting representations in Lemma 1. The asymptotic distribution of  $t_{\phi_{ML}=1}$  is a scaled mixture of a Dickey-Fuller and a standard normal random variables with a mixing coefficient that depends on the degree of conditional heteroskedasticity and non-normality of the errors. In the case of normally distributed errors ( $K = F$ ), the Dickey-Fuller distribution provides an upper bound for the critical values of  $t_{\phi_{ML}=1}$ . As the degree of conditional heteroskedasticity increases,<sup>1</sup> more weight is assigned to the standard normal distribution and the corresponding smaller critical values increase the power of the test. The limiting representations for models with constant and linear trend can be obtained by replacing  $B_1(r)$  in (3) and (4) by its demeaned  $B_1(r) - \int_0^1 B_1(s) ds$  and detrended  $B_1(r) - \int_0^1 (4 - 6s) B_1(s) ds - r \int_0^1 (12s - 6) B_1(s) ds$  counterparts, respectively.

Another version of the test standardizes  $(\hat{\phi}_{ML} - 1)$  with the robust variance covariance matrix (Bollerslev and Wooldridge, 1992)  $\left[-\sum_{t=1}^T \frac{\partial^2 l_t(\phi, \delta)}{\partial \phi^2}\right]^{-1} \left[-\sum_{t=1}^T \left(\frac{\partial l_t(\phi, \delta)}{\partial \phi}\right)^2\right] \left[-\sum_{t=1}^T \frac{\partial^2 l_t(\phi, \delta)}{\partial \phi^2}\right]^{-1}$ , evaluated at the ML estimates of  $\phi$  and  $\delta$ , whose limiting distribution is given by  $\rho \frac{\int_0^1 B_1(r) dB_1(r)}{\left(\int_0^1 B_1^2(r) dr\right)^{1/2}} +$

---

<sup>1</sup>Boswijk (2001) derives an approximate expression of  $\rho$  in terms of the GARCH parameters as  $\rho \approx \sqrt{\frac{(1-\alpha-\beta)(1-\beta^2)}{(1-\alpha-\beta+\alpha^2)(1-\beta^2+2\alpha^2)}}$ . It is then easy to see that high persistence in the conditional variance ( $\alpha + \beta$  near one) is typically associated with low values of  $\rho$ .

$\sqrt{1 - \rho^2}\xi$ . This test is expected to have more robust size properties with possibly non-normally distributed errors although at the cost of moderate power losses for Gaussian errors.

Despite its potential for non-trivial power improvements, the test in (4) has the unappealing property that its asymptotic distribution is non-pivotal and depends on nuisance parameters. In principle, one could tabulate critical values for the test  $t_{\phi_{ML}=1}\sqrt{\widehat{F}/\widehat{K}}$  on a grid of values for  $\rho$  (Seo, 1999), where the nuisance parameters are estimated from the data, although this makes the testing procedure somewhat cumbersome. More importantly, the nuisance parameters involve infinite sums and estimates of  $\alpha$ ,  $\beta$ ,  $\kappa$  and  $h$  that enter in a highly nonlinear fashion which could impair the precision with which these quantities are computed. As we demonstrate below, this may lead to severe size distortions of the tests even for large sample sizes. The bootstrap method that we propose in this paper proves to be very useful for approximating the finite-sample distribution of  $t_{\phi_{ML}=1}$  as it avoids the explicit calculation of the nuisance parameters. In addition to the substantially improved size properties of the unit root test, the straightforward implementation of the bootstrap offers practical advantages and can be easily extended to processes that accommodate more general serial correlation and conditional heteroskedasticity structure.

### 3 Bootstrap Approximation

In this section, we propose a bootstrap method for approximating the finite-sample distribution of the unit root test  $t_{\phi_{ML}=1}$ . We start by discussing the bootstrap procedures based on resampling the symmetrized residuals and generating repeated samples under the null of a unit root. In proving the asymptotic validity of the bootstrap, we first verify that the bootstrap samples satisfy the conditions of Assumption 1 and the effect of the initial conditions is asymptotically negligible. Then, we develop a bootstrap invariance principle with conditionally heteroskedastic errors and establish the weak convergence of the bootstrap statistic to the limiting distribution in Lemma 1.

#### 3.1 Description of the Bootstrap Procedure

Let  $\{y_1, y_2, \dots, y_T\}$  be a sequence of  $T$  observations generated by model (1). As argued above, the conditional mean and variance parameters of (1) can be estimated separately. Let  $\widehat{\delta} = (\widehat{\omega}, \widehat{\alpha}, \widehat{\beta})'$  denote the ML estimates of the GARCH parameters,  $\{\widehat{h}_t\}$  be the conditional variance computed



recursively from these estimates for some initial value  $h_0$  and  $\hat{\phi}_{ML}$  denote the one-step or iterated MLE of  $\phi$  introduced in the previous section.

Define the residuals  $\tilde{\varepsilon}_t = y_t - \hat{\phi}_{ML} y_{t-1}$ . While these residuals could also be constructed imposing the null of a unit root ( $\phi = 1$ ), we follow Paparoditis and Politis (2003) and compute the residuals using the MLE of  $\phi$  which helps to retain the important characteristics of the data and improve the power of the unit root test. We then construct the recentered standardized residuals as  $\tilde{\eta}_t = \tilde{\varepsilon}_t / \sqrt{\hat{h}_t} - T^{-1} \sum_{i=1}^T \tilde{\varepsilon}_i / \sqrt{\hat{h}_i}$  for  $t = 1, 2, \dots, T$  with empirical distribution function denoted by  $\tilde{F}_T(\eta) = T^{-1} \sum_{t=1}^T I(\tilde{\eta}_t \leq \eta)$  that is used for resampling. Since the underlying distribution of  $\eta_t$  is assumed to be symmetric (Assumption 1, part (a)), we need to ensure that the empirical distribution from which the bootstrap samples are drawn is also symmetric. For this reason, we construct the collection  $\{\pm \tilde{\eta}_1, \pm \tilde{\eta}_2, \dots, \pm \tilde{\eta}_T\}$  which is symmetric by construction (Jing, 1995).

The bootstrap procedure for approximating the distribution of  $t_{\phi_{ML}=1}$  takes the following steps. First, draw a random sample  $\{\eta_1^*, \eta_2^*, \dots, \eta_T^*\}$  from  $\{\pm \tilde{\eta}_1, \pm \tilde{\eta}_2, \dots, \pm \tilde{\eta}_T\}$  with replacement and for initial conditions  $h_0^*$  and  $y_0^*$ , construct a bootstrap sample recursively as

$$\begin{aligned} h_t^* &= \hat{\omega} + (\hat{\beta} + \hat{\alpha} \eta_{t-1}^{*2}) h_{t-1}^* \\ y_t^* &= y_{t-1}^* + \sqrt{h_t^*} \eta_t^*. \end{aligned}$$

The bootstrap sample  $\{y_1^*, y_2^*, \dots, y_T^*\}$  is first used to get the bootstrap QMLE estimates  $\delta^* = (\omega^*, \alpha^*, \beta^*)$  from  $\varepsilon_t^* = y_t^* - \phi_{LS}^* y_{t-1}^*$ , where  $\phi_{LS}^* = (\sum_{t=1}^T y_{t-1}^{*2})^{-1} (\sum_{t=1}^T y_t^* y_{t-1}^*)$ . Then, the one-step bootstrap QMLE of  $\phi$  is obtained as

$$\phi_{ML}^* = \tilde{\phi}^* - \left[ \sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta^*)}{\partial \phi^2} \right]_{\phi=\tilde{\phi}^*}^{-1} \left[ \sum_{t=1}^T \frac{\partial l_t^*(\phi, \delta^*)}{\partial \phi} \right]_{\phi=\tilde{\phi}^*},$$

where  $\tilde{\phi}^*$  is a preliminary consistent estimate, typically  $\phi_{LS}^*$ . The iterative bootstrap estimator can be computed by updating the estimates of  $\delta^*$  and  $\phi_{ML}^*$  until convergence. The estimators  $\delta^*$  and  $\phi_{ML}^*$  are finally used to calculate the Hessian  $\left[ -\sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta)}{\partial \phi^2} \right]_{\phi=\phi_{ML}^*, \delta=\delta^*}^{-1}$  and the  $t$ -statistic of a unit root  $t_{\phi_{ML}=1}^* = \left[ -\sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta)}{\partial \phi^2} \right]_{\phi=\phi_{ML}^*, \delta=\delta^*}^{1/2} (\phi_{ML}^* - 1)$ .

This algorithm is repeated  $B$  times<sup>2</sup> and each time the bootstrap unit root statistic  $t_{\phi_{ML}=1}^*$  is computed. Let  $P^*$  denote the distribution of  $(y_1^*, y_2^*, \dots, y_T^*)$  conditional on the sample  $(y_1, y_2, \dots, y_T)$

---

<sup>2</sup>See, for example, Davidson and MacKinnon (2000) for guidance in selecting the number of bootstrap replications.

and  $G_T^*(x) = P^*(t_{\phi_{ML}=1}^* \leq x)$  be the bootstrap distribution of  $t_{\phi_{ML}=1}^*$ . Bootstrap critical values can be obtained by taking the corresponding quantile of  $G_T^*(x)$  and bootstrap  $p$ -values of the unit root test are constructed as  $B^{-1} \sum_{j=1}^B I(t_{\phi_{ML}=1}^* \leq t_{\phi_{ML}=1})$ . The procedure can be adapted easily to models with deterministic components.

### 3.2 Asymptotic Validity of the Bootstrap

This section analyzes the asymptotic properties of the symmetrized-residual bootstrap procedure. We first demonstrate that the bootstrap samples satisfy the conditions of Assumption 1. We also show that the initial values used for generating bootstrap samples do not affect the asymptotic distribution of the test statistic. We then establish the bootstrap invariance principle for partial sums of processes with GARCH errors and prove the weak convergence of the bootstrap unit root test statistic to the asymptotic distribution (4) in Lemma 1.

From the properties of the MLE estimator  $\hat{\delta}$  and the constraints imposed in the estimation of the GARCH parameters, it is easy to verify that part (b) of Assumption 1 still holds for the bootstrap data generating process. As a result, we focus on establishing if the bootstrap samples satisfy the conditions of parts (a) and (c) of Assumption 1.

Let  $d_2(\cdot)$  denote the Mallows metric<sup>3</sup> of degree 2, defined as  $d_2(F_X, F_Z) = \inf \left( E |X - Z|^2 \right)^{1/2}$  over all joint distributions for the random variables  $X$  and  $Z$  with marginal distributions  $F_X$  and  $F_Z$ . Also, let  $\tilde{F}_T^{sym}(\eta) = (2T)^{-1} \sum_{t=1}^T [I(\tilde{\eta}_t \leq \eta) + I(-\tilde{\eta}_t \leq \eta)]$  denote the empirical distribution function of the symmetrized recentered residuals  $\{\pm\tilde{\eta}_1, \pm\tilde{\eta}_2, \dots, \pm\tilde{\eta}_T\}$  and  $F$  be the true distribution of the standardized errors  $\eta_t$ . We use the Mallows metric  $d_2$  to show that the symmetrized empirical distribution function of the recentered standardized residuals provides a good approximation to the true distribution function and the bootstrap errors satisfy the conditions for establishing the bootstrap invariance principle.

LEMMA 2. *Let  $E^*$  and  $Var^*$  refer to the expected value and variance of  $P^*$ ,  $\{\eta_t^*\}_{t=1}^T$  be drawn with replacement from  $\tilde{F}_T^{sym}(\eta)$  and suppose that Assumption 1 holds. Then, (a)  $d_2(\tilde{F}_T^{sym}, F) \rightarrow 0$  as  $T \rightarrow \infty$ , (b)  $E(\varepsilon_t^*) = 0$ , (c)  $Var^*(\varepsilon_t^*) = \sigma^2$  as  $T \rightarrow \infty$ , and (d)  $E^*(\varepsilon_t^*)^3 = 0$ .*

PROOF. See Appendix A.2.

---

<sup>3</sup>For the properties of the Mallows metric, see Section 8 in Bickel and Friedman (1981).

The bootstrap sequences  $\{h_t^*\}$  and  $\{\varepsilon_t^*\}$  are constructed for some initial values  $h_0^*$  and  $\eta_0^*$ . Auxiliary Lemma 2 in Appendix A.1 establishes that if  $\eta_0^*$  is drawn from  $\tilde{F}_T^{sym}(\eta)$  and  $h_t^*$  is initialized from its invariant measure, the bootstrap sequences  $\{h_t^*\}$  and  $\{\varepsilon_t^*\}$  are strictly stationary and ergodic. Furthermore, Auxiliary Lemma 3 in Appendix A.1 shows that the expected difference (under  $P^*$ ) of partial sums constructed from sequences that start from infinite past and finite past tend to zero as  $T \rightarrow \infty$ .

The following lemma demonstrates that different initial values of  $h_t^*$  have no asymptotic effect on the bootstrap procedure.

LEMMA 3. *Define the processes  $\xi_t^* = \lambda_1 \varepsilon_t^* + \lambda_2 [\frac{\varepsilon_t^*}{h_t^*} - \frac{\hat{\alpha}}{h_t^*} (\frac{\varepsilon_t^{*2}}{h_t^*} - 1) \sum_{k=1}^{t-1} \hat{\beta}^{k-1} \varepsilon_{t-k}^*]$  and  $S_{[Tr]} = T^{-1/2} \sum_{t=1}^{[Tr]} \xi_t^*$  for  $\{0 \leq r \leq 1\}$ , where  $\lambda = (\lambda_1, \lambda_2)'$  is a constant vector with  $\lambda \lambda' \neq 0$ . Let  $h_{01}^*$  and  $h_{02}^*$  are two different initial values of  $h_t^*$  and  $(h_{t1}^*, h_{t2}^*)$ ,  $(\varepsilon_{t1}^*, \varepsilon_{t2}^*)$  and  $(\xi_{t1}^*, \xi_{t2}^*)$  are bootstrap sequences corresponding to these initial values, respectively.*

*Then, under Assumption 1 and as  $T \rightarrow \infty$ , (a)  $E^* |h_{t1}^* - h_{t2}^*| \rightarrow 0$ , (b)  $E^* |\varepsilon_{t1}^* - \varepsilon_{t2}^*| \rightarrow 0$ , (c)  $E^* \left| \frac{1}{\sqrt{T}} \sum_{i=1}^T \varepsilon_{i1}^* - \frac{1}{\sqrt{T}} \sum_{i=1}^T \varepsilon_{i2}^* \right| = O(T^{-1/2})$ , and (d)  $E^* \left| S_{[Tr]}^{(1)} - S_{[Tr]}^{(2)} \right| = O(T^{-1/2})$ , where  $S_{[Tr]}^{(1)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \xi_{t1}^*$  and  $S_{[Tr]}^{(2)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \xi_{t2}^*$ .*

PROOF. See Appendix A.2.

Finally, we show that the bootstrap delivers consistent estimates of the nuisance parameters that enter the limiting distribution of the unit root test.

LEMMA 4. *Under Assumption 1 and as  $T \rightarrow \infty$ , (a)  $E^*(h_t^*) \rightarrow E(h_t)$ , (b)  $E^*(1/h_t^*) \rightarrow E(1/h_t)$ , (c)  $K^* \rightarrow K$ , and (d)  $F^* \rightarrow F$ .*

PROOF. See Appendix A.2.

Now we can establish the bootstrap invariance principle for partial sums of GARCH processes.

LEMMA 5. *Under Assumption 1,*

$$\left( T^{-1/2} \sum_{t=1}^{[Tr]} \varepsilon_t^*, T^{-1/2} \sum_{t=1}^{[Tr]} \left[ \frac{\varepsilon_t^*}{h_t^*} + (1 - \eta_t^{*2}) \frac{\hat{\alpha}}{h_t^*} \sum_{j=1}^t \hat{\beta}^{j-1} \varepsilon_{t-j}^* \right] \right) \Rightarrow [W_1(r), W_2(r)]$$

for all  $r \in [0, 1]$ , conditionally on the sample  $(y_1, y_2, \dots, y_T)$ , where  $[W_1(r), W_2(r)]$  is a bivariate Brownian motion in  $D[0, 1] \times D[0, 1]$  with mean zero and covariance matrix  $\Omega = r \begin{pmatrix} E(h_t) & 1 \\ 1 & K \end{pmatrix}$ , where  $K$  is defined in Lemma 1.

PROOF. See Appendix A.2.

The results in Lemmas 2 to 5 provide sufficient conditions for the asymptotic validity of the bootstrap procedure. The next theorem shows that the bootstrap approximation to the distribution of the  $t_{\phi_{ML}=1}$  test converges weakly to the limiting distribution in Lemma 1 which implies that the bootstrap is first-order asymptotically correct.

THEOREM 1. Under Assumption 1 and the null hypothesis  $H_0 : \phi = 1$ , for any  $x \in \mathbb{R}$  and  $\epsilon > 0$ ,

$$\lim_{T \rightarrow \infty} \Pr \left\{ \sup_x \left| P^*(t_{\phi_{ML}=1}^* \leq x) - P(t_{\phi_{ML}=1} \leq x) \right| > \epsilon \right\} = 0,$$

where  $P(t_{\phi_{ML}=1} \leq x)$  is the limiting distribution (4) of the  $t_{\phi_{ML}=1}$  test in Lemma 1.

PROOF. See Appendix A.2.

Theorem 1 implies that the critical and  $p$ -values for the unit root test with GARCH errors can be approximated by the proposed bootstrap that avoids the explicit estimation of nuisance parameters. An interesting extension that is beyond of the scope of this paper is to study the power of the bootstrap test under the alternative and show that it converges to the power function of the asymptotic test as in Swensen (2003). Also, while investigating the higher-order accuracy of the bootstrap might be interesting, the bootstrap is not expected to offer any asymptotic refinements since the test statistic is not pivotal.

The next section shows that the asymptotic distribution (4) provides a very poor approximation to the finite-sample distribution of the unit root test when the degree of conditional heteroskedasticity is high. This seems to arise from the imprecise estimation of the nuisance parameters as the conditional heteroskedasticity is close to an integrated GARCH process. In contrast, the size of the bootstrap-based test is near the nominal level across all GARCH parameterizations without any adverse effects on the power.

## 4 Numerical Illustrations

### 4.1 Monte Carlo Simulation

This section reports the results from a Monte Carlo experiment that assesses the size and power properties of the asymptotic and bootstrap unit root tests in models with GARCH errors. Repeated sample paths are generated from the following model

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &= \sqrt{h_t} \eta_t \\ h_t &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}, \end{aligned} \tag{5}$$

where  $\eta_t \sim iid(0, 1)$ . We consider three error distributions: Gaussian,  $t$  distribution with 7 degrees of freedom and chi-square distribution with 5 degrees of freedom that are appropriately standardized to have mean zero and variance one. The sample sizes are  $T = 200$  and  $400$  and the number of Monte Carlo replications is 2,000.

The autoregressive parameter  $\phi$  takes values of 1 and 0.92 in evaluating the size and the power of the unit root test, respectively. We also normalize the unconditional variance to be one by setting  $\omega = 1 - \alpha - \beta$ . The performance of the tests is evaluated for different degrees of conditional heteroskedasticity that cover the conditional homoskedastic case ( $\alpha + \beta = 0$ ) and some highly persistent GARCH specifications ( $\alpha + \beta = 0.999$ ). We consider specifications that are typically estimated from financial data (for example,  $(\alpha = 0.399, \beta = 0.6)$  and  $(\alpha = 0.199, \beta = 0.8)$ ) as well as specifications (large  $\alpha$  and small  $\beta$ , for instance) that are not frequently encountered in economic applications. It should be noted that while all specifications satisfy the moment condition  $E\varepsilon_t^2 < \infty$ , most of the considered GARCH parameterizations render  $E\varepsilon_t^4$  infinite.

We investigate the empirical size and power performance of the asymptotic test based on the OLS estimator ( $ASY - DF$ ), the DF test with critical values approximated by the wild bootstrap ( $BOOT - DF$ ), the asymptotic test based on the iterated ML estimator of the GARCH model ( $ASY - GARCH$ ) and its bootstrap analog ( $BOOT - GARCH$ ) discussed in Section 3. All tests are constructed using demeaned data which is equivalent to including an intercept in the estimated models. In the ML estimation of the GARCH parameters, we impose the restriction  $\alpha + \beta < 1$ .

The GARCH bootstrap generates samples under the null of a unit root by resampling the centered, symmetrized standardized residuals. These samples are used to approximate the distribution of the unit root test with 199 bootstrap replications that delivers the corresponding bootstrap critical values. The asymptotic critical values for the test based on the OLS estimator are obtained from the Dickey-Fuller tables. For the asymptotic test based on the ML estimator with GARCH errors, we use the true values of  $\alpha$ ,  $\beta$  and  $\kappa$  to obtain the values of the nuisance parameters  $F$ ,  $K$  and  $\rho$  (by truncating the infinite sums at a large integer value) and then interpolate the appropriate critical values from Table 3 in Seo (1999).

The empirical rejection probabilities under the null of a unit root at 1%, 5% and 10% nominal levels for standard normal errors and sample sizes 200 and 400 are reported in Tables 1 and 2. The asymptotic DF test is well sized in the conditionally homoskedastic case and slightly overrejects for low to moderate degrees of conditional heteroskedasticity. As the GARCH persistence approaches the unit boundary, the size distortions of the DF test are substantial (see also Valkanov, 2005) and are bigger when  $\alpha$  exceeds  $\beta$ . Several recent papers (Beare, 2008; Cavaliere and Taylor, 2008, 2009) have proposed modified unit root test procedures that are robust to the presence of certain types of conditional heteroskedasticity.<sup>4</sup> Here, we consider the wild bootstrap approach of Cavaliere and Taylor (2008) who extend the results of Gonçalves and Kilian (2004, 2007) to unit root models with nonstationary volatility. The second column of Tables 1 and 2 presents the results based on the wild bootstrap method. The wild bootstrap reduces the size distortions of the asymptotic DF test but there are still some relative large overrejections when the sum of the GARCH parameters is near unity. This reflects the stronger moment requirements on the errors that are needed for establishing the validity of the wild bootstrap (Cavaliere and Taylor, 2008).

The results for the *ASY – GARCH* test  $t_{\phi_{ML}=1}$  are reported in the third column of Tables 1 and 2. While the size distortions of this test are smaller than those of the DF test, they are still fairly large despite the fact that the *ASY – GARCH* test is designed to handle explicitly the presence

---

<sup>4</sup>Some other popular methods for size correction may not be valid or appropriate in our context. For example, using a robust variance covariance matrix tends to reduce the size distortions (Kim and Schmidt, 1993) but the consistency of this procedure for nonstationary processes has not been formally established. Also, while the resampling scheme that incorporates the GARCH structure of the model can certainly be used for the DF test, it is not obvious why one would employ it for this test and not for the more powerful test based on the ML estimator.

of conditional heteroskedasticity. Substantial overrejections occur when the GARCH specification borders an integrated GARCH process.<sup>5</sup>

In contrast to the large size distortions of the asymptotic tests, our proposed bootstrap method (last column in Tables 1 and 2) controls the size of the unit root test with GARCH errors uniformly across all GARCH specifications and nominal levels. This impressive performance of the bootstrap unit root test is achieved despite the small number of bootstrap replications. Overall, our bootstrap procedure proves to be very effective for correcting the overrejections of the *ASY – GARCH* test.

Tables 3 and 4 report the empirical power of the unit root tests with simulated data from model (5) with  $\phi = 0.92$ ,  $\eta_t \sim N(0, 1)$  and  $T = 200$  and  $400$ . The rejection probabilities for the asymptotic tests (*ASY – DF* and *ASY – GARCH*) is size-adjusted power whereas the power of the bootstrap tests (*BOOT – DF* and *BOOT – GARCH*) is raw power. One interesting observation that emerges from the results is that the asymptotic DF test is not able to detect any deviations from the null hypothesis when the conditional heteroskedasticity is very strong and  $T = 200$ . For example, if  $(\alpha = 0.6, \beta = 0.399)$  and  $(\alpha = 0.8, \beta = 0.199)$ , the size-adjusted power of the DF test is only 6.70% and 7.12% at 10% nominal level and  $T = 200$ . Even for the parameterization  $(\alpha = 0.399, \beta = 0.6)$  that is more often encountered in financial applications, the power is 9.55% at 10% nominal level. As the sample size gets larger,<sup>6</sup> the power of the asymptotic DF test improves but it is still well below that of the *ASY – GARCH* test. Interestingly, the wild bootstrap has better power than its asymptotic analog although some of the power gains are due to the overrejections under the null of a unit root reported in Tables 1 and 2.

The tests that incorporate the GARCH structure of the model suffer only a small power loss in the conditionally homoskedastic case but offer moderate to very large power gains when the degree of conditional heteroskedasticity in the GARCH specification increases. These substantial power

---

<sup>5</sup>Our numerical experiments suggest that these overrejections are due to imprecise estimation of the nuisance parameters  $\sigma^2 = E(h_t)$ ,  $E(1/h_t)$  and  $K$  as  $\alpha + \beta$  is close to one. For example, when  $\alpha + \beta = 0.99$ , the estimates of  $\sigma^2$  start to deviate significantly from 1 and tend to be biased towards 0. The difference becomes even more extreme for  $\alpha + \beta = 0.999$  and large values of  $\alpha$  (Gospodinov and Tao, 2009).

<sup>6</sup>All computations are performed in GAUSS. The computational time increases roughly 1.5 times when the sample size doubles from 200 to 400. More precisely, the average time with  $T = 200$  is 2.2 seconds per Monte Carlo replication while with  $T = 400$ , it is 3.3 seconds (2.66GHz Intel Core 2 processor).

improvements, combined with the size correction property of the bootstrap method, illustrate the potential of the ML-based tests to detect the mean reversion in processes with strong conditional heteroskedasticity. The raw power of the bootstrap test is very close, albeit slightly below, the (typically infeasible in practice) size-adjusted power of *ASY – GARCH*. Davidson and MacKinnon (2006) analyze the discrepancy that arises between the rejection probabilities of the bootstrap test and the size-adjusted power of the asymptotic test and suggest possible ways of minimizing it.

We now turn our attention to the size and power properties of the unit root tests with non-normal errors. The  $t$  distribution with 7 degrees of freedom is often used in econometric applications to capture the fatter tails of financial data at weekly or monthly frequency. The  $\chi^2$  distribution does not satisfy the symmetry condition in part (a) of Assumption 1 and is used to investigate the sensitivity of the tests to asymmetric errors. While the  $\chi^2$  distribution with 5 degrees of freedom produces much larger asymmetry than that typically observed in economic and financial data, it would be interesting to assess the behavior of the tests for more extreme specifications. In this case, we do not symmetrize the residuals as in Section 3.1 and allow the bootstrap procedure to adapt to the shape of the estimated error distribution. To preserve space, we only report the empirical size and power of the tests for  $T = 200$ . The results for non-normal errors (Tables 5 and 6 for  $t_7$  distribution and Tables 7 and 8 for  $\chi_5^2$  distribution) can be summarized as follows.

The empirical size and power of the asymptotic DF test appears to be fairly similar across the different error distributions. The bootstrap DF test tends to overreject more for non-normal errors, especially then the sum of the GARCH parameters is near unity (in some cases, the empirical size is close to and above 20% at 10% nominal level). Overall, the wild bootstrap appears to be an effective tool for reducing the large size distortions of the asymptotic DF test although, strictly speaking, it is not theoretically valid for most of the GARCH specifications considered in this paper. The performance of the *ASY – GARCH* test deteriorates further in the case of  $\chi_5^2$ -distributed errors as predicted by theory. While the *BOOT – GARCH* test also tends to overreject for asymmetric errors (by 2-3 percentage points at 10% nominal level), it is still very well sized and delivers significant power improvements.



## 4.2 Testing for a Unit Root in U.S. Interest Rates

The correct specification of the dynamics of interest rates plays an important role in derivative pricing, hedging and term structure modeling. For example, most diffusion models of spot interest rate that are used for bond valuation impose a mean reverting behavior on the underlying process. Yet, unit root tests for post-war U.S. interest rates rarely reject the null of a unit root which requires that this nonstationarity is taken into account in modeling and long-run forecasting of interest rates. This empirical finding not only creates some tension between the dynamics of interest rates in theoretical finance and the specification adopted in practice but it may also cause substantial size distortions in testing the parameters in term structure models (Elliott, 1998).

While the conditional heteroskedasticity is a widely documented characteristic of interest rates, the unit root tests typically do not incorporate explicitly the strong GARCH effect into the testing procedure. We re-examine the possibility of a mean reversion in U.S. interest rates using the bootstrap test proposed in this paper. The data employed in the analysis include the Federal Funds rate, 3-month Treasury bill rate (secondary market), 1-, 5- and 10-year Treasury bond yields (constant maturity) and the default premium constructed as the difference between the Aaa and Baa corporate bond yields. The series are annualized rates at monthly frequency covering the period July 1954 - November 2008 and are downloaded from Table H.15 of the Federal Reserve Statistical Release (<http://www.federalreserve.gov/releases/h15/data.htm>). The dynamics of the five interest rates and the default premium are plotted in Figures 1 and 2, respectively. The graphs show that all series are characterized by high persistence over the sample period. The short-term interest rates appear to be more volatile than the long-term rates and the dynamics become smoother as the time to maturity increases. Finally, the sum of the estimated GARCH parameters for all interest rates is very close to one which indicates a strong volatility clustering.

The results from the Dickey-Fuller and the GARCH-based unit root tests are reported in Table 9. Since the interest rates do not exhibit any trending behavior, we consider a model that includes an intercept but not a linear trend. The values of the DF statistic for all interest rate processes do not exceed the asymptotic critical values at 5% and 10% significance level (-2.86 and -2.57, respectively). The asymptotic  $p$ -values of the DF tests are between 0.27 and 0.62 and provide no evidence against the null of a unit root. This appears to be due to the low power of the DF test

for detecting mean reversion in processes with strong conditional heteroskedasticity reported in our simulation study. The results from our bootstrap test with GARCH errors stand in sharp contrast with this finding. The bootstrap  $p$ -values of the *BOOT – GARCH* test suggest that the null of a unit root can be rejected at 5% significance level for all interest rates except for the 10-year yield whose bootstrap  $p$ -value is 0.068. Incorporating the GARCH structure of interest rates into the testing procedure confirms the substantial power gains documented in the previous section. This rejection of the unit root hypothesis also lends empirical support to the mean reverting diffusion specification that is typically used in financial economics to describe the dynamics of short-term interest rates.

## 5 Conclusion

This paper proposes a bootstrap test for a unit root in processes with GARCH errors and shows its asymptotic validity under very weak moment and distributional assumptions. The proposed method offers several important advantages over its asymptotic counterpart and the existing tests that do not exploit the information in the conditional variance. First, the test delivers impressive power gains by explicitly incorporating the GARCH structure of the errors, especially for highly persistent GARCH specifications with power improvements over the DF-type tests. While the asymptotic counterpart of the test requires the computation of nuisance parameters and suffers from relatively large size distortions, the proposed bootstrap procedure is straightforward to implement and appears to control the size uniformly over all possible GARCH specifications that guarantee the existence of second moments of the errors. Finally, while generalizing the asymptotic theory to more complicated setups would be quite involved, our bootstrap method can be easily adapted to models with a lag length that goes to infinity at certain rate, asymmetric errors and different types of conditional heteroskedasticity (other models from the GARCH class, stochastic volatility models etc.)

## References

- [1] Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H., Taylor, R.L. (1991). Bootstrapping unstable first-order autoregressive processes. *Annals of Statistics* 19:1098-1101.
- [2] Basawa, I.V., Mallik, A.K., McCormick, W.P., Taylor, R.L. (1989). Bootstrapping explosive autoregressive processes. *Annals of Statistics* 17:1479-1486.
- [3] Beare, B.K. (2008). Unit root testing with unstable volatility. Nuffield College Economics Working Paper No. 2008-06.
- [4] Bickel, P.J., Freedman, D. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics* 9:1196-1217.
- [5] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 31:307-328.
- [6] Bollerslev, T., Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews* 11:143-172.
- [7] Boswijk, H.P. (2001). Testing for a unit root with near-integrated volatility. Unpublished manuscript, Universiteit van Amsterdam.
- [8] Carrasco, M., Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18:17-39.
- [9] Cavaliere, G., Taylor, A.M.R. (2008). Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory* 24:43-71.
- [10] Cavaliere, G., Taylor, A.M.R. (2009). Bootstrap M unit root tests. *Econometric Reviews* 28:393-421.
- [11] Davidson, R., MacKinnon, J.G. (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews* 19:55-68.
- [12] Davidson, R., MacKinnon, J.G. (2006). The power of bootstrap and asymptotic tests. *Journal of Econometrics* 133:421-441.
- [13] Elliott, G. (1998). On the robustness of cointegration methods when regressors almost have unit roots. *Econometrica* 66:149-158.
- [14] Ferretti, N., Romo, J. (1996). Unit root bootstrap tests for AR(1) models. *Biometrika* 83:849-860.
- [15] Francq, C., Zakořan, J.M. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22:815-834.
- [16] Gonçalves, S., Killan, L. (2004). Bootstrapping autoregression with conditional heteroskedasticity of unknown form. *Journal of Econometrics* 123:89-120.
- [17] Gonçalves, S., Killan, L. (2007). Asymptotic and bootstrap inference for AR(infinity) processes with conditional heteroskedasticity. *Econometric Reviews* 26:609-641.
- [18] Gospodinov, N. (2008). Asymptotic and bootstrap tests for linearity in a TAR-GARCH(1,1) model with a unit root. *Journal of Econometrics* 146:146-161.
- [19] Gospodinov, N., Tao, Y. (2009). Bootstrap unit root tests in models with GARCH(1,1) errors. Working paper 09-001, Concordia University.

- [20] Heimann, G., Kreiss J.P. (1996). Bootstrapping general first order autoregression. *Statistics and Probability Letters* 30:87-98.
- [21] Jing, B.Y. (1995). Some resampling procedures under symmetry. *Australian Journal of Statistics* 37:337-344.
- [22] Kim, K., Schmidt, P. (1993). Unit root tests with conditional heteroskedasticity. *Journal of Econometrics* 59:287-300.
- [23] Lee, S.W., Hansen, B.E. (1994). Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* 10:29-52.
- [24] Ling, S., Li, W.K. (1998). Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general autoregressive heteroskedastic errors. *Annals of Statistics* 26:84-125.
- [25] Ling, S., Li, W.K. (2003). Asymptotic inference for unit root processes with GARCH (1,1) errors. *Econometric Theory* 19:541-564.
- [26] Ling, S., Li, W.K., McAleer, M. (2003). Estimation and testing for unit root processes with GARCH (1,1) errors: Theory and Monte Carlo evidence. *Econometric Reviews* 22:179-202.
- [27] MacKinnon, J.G. (1996). Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics* 11:601-618.
- [28] Nelson, D.B. (1990). Stationarity and persistence in the GARCH (1,1,) model. *Econometric Theory* 6:318-334.
- [29] Paparoditis, E., Politis, D.N. (2003). Residual-based block bootstrap for unit root testing. *Econometrica* 71:813-855.
- [30] Park, J.Y. (2003). Bootstrap Unit Root Test. *Econometrica* 71:1845-1895
- [31] Pascual, L., Romo, J., Ruiz, E. (2000). Forecasting returns and volatilities in GARCH process using the bootstrap. Working paper 00-68, Universidad Carlos III de Madrid.
- [32] Seo, B. (1999). Distribution theory for unit root tests with conditional heteroskedasticity. *Journal of Econometrics* 91:113-144.
- [33] Swensen, A.R. (2003). A note on the power of bootstrap unit root tests. *Econometric Theory* 19:32-48.
- [34] Valkanov, R. (2005). Functional central limit theorem approximations and the distribution of the Dickey-Fuller test with strongly heteroskedastic data. *Economics Letters* 86:427-433.

## A Appendix: Auxiliary Lemmas and Proofs

### A.1 Auxiliary Lemmas

AUXILIARY LEMMA 1. Under Assumption 1, (a)  $\widehat{h}_t - h_t = o_p(1) + O(\beta^t)$  and (b)  $\frac{\varepsilon_t^2}{h_t \widehat{h}_t} = O_p(1)$ .

PROOF. For proof of part (a), see Gospodinov (2008). For part (b) note that  $\frac{\varepsilon_t^2}{h_t \widehat{h}_t} < \frac{\varepsilon_t^2}{\omega \widehat{\omega}} \leq \frac{\varepsilon_t^2}{\gamma^2}$ , where  $\gamma = \min\{\omega, \widehat{\omega}\} > 0$ . Since  $E\left(\frac{\varepsilon_t^2}{\gamma^2}\right) = \frac{\sigma^2}{\gamma^2}$  and  $\frac{\varepsilon_t^2}{h_t \widehat{h}_t} \geq 0$ , it is easy to show that  $\frac{\varepsilon_t^2}{h_t \widehat{h}_t} = O_p(1)$ . ■

AUXILIARY LEMMA 2. Let  $h_t^* = \widehat{\omega} \left[1 + \sum_{k=1}^{\infty} \Pi_{i=1}^k \left(\widehat{\alpha} \eta_{t-i}^{*2} + \widehat{\beta}\right)\right]$  and  $\varepsilon_t^* = \eta_t^* \sqrt{\widehat{\omega} \left[1 + \sum_{k=1}^{\infty} \Pi_{i=1}^k \left(\widehat{\alpha} \eta_{t-i}^{*2} + \widehat{\beta}\right)\right]}$  and suppose that  $\eta_0^*$  is drawn from  $\widehat{F}_T^{sym}(\eta)$  and the sequence  $\{h_t^*\}$  is initialized from its invariant measure. Then,  $\{h_t^*\}$  and  $\{\varepsilon_t^*\}$  are strictly stationary and ergodic processes.

PROOF. The proof follows directly from Theorem 2 in Nelson (1990). ■

AUXILIARY LEMMA 3. Let  $\xi_t = \lambda_1 \varepsilon_t^* + \lambda_2 \left[\frac{\varepsilon_t^*}{h_t^*} - \frac{\widehat{\alpha}}{h_t^*} (\eta_t^{*2} - 1) \sum_{k=1}^{t-1} \widehat{\beta}^{k-1} \varepsilon_{t-k}^*\right]$  and  $\xi_t' = \lambda_1 \varepsilon_t^* + \lambda_2 \left[\frac{\varepsilon_t^*}{h_t^*} - \frac{\widehat{\alpha}}{h_t^*} (\eta_t^{*2} - 1) \sum_{k=1}^{\infty} \widehat{\beta}^{k-1} \varepsilon_{t-k}^*\right]$ , and denote  $S_{[Tr]} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \xi_t$  and  $S'_{[Tr]} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} \xi_t'$  for  $\{0 \leq r \leq 1\}$ . Then,  $E^* \left| S_{[Tr]} - S'_{[Tr]} \right| \rightarrow 0$ .

PROOF. The proof is similar to that of (4.6) in Lemma 4.2 in Ling and Li (2003). More specifically,

$$\begin{aligned} E^* \left| S_{[Tr]} - S'_{[Tr]} \right| &\leq \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} E^* \left| \frac{\lambda_2}{h_t^*} \left( \frac{\varepsilon_t^{*2}}{h_t^*} - 1 \right) \left( \widehat{\alpha} \sum_{k=t}^{\infty} \widehat{\beta}^{k-1} \varepsilon_{t-k}^* \right) \right| \\ &\leq \frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} E^* \left| \frac{\lambda_2}{\sqrt{\widehat{\alpha}_0}} (\eta_t^{*2} - 1) \right| E^* \left| \widehat{\alpha} \sum_{k=t}^{\infty} \widehat{\beta}^{k-1} \frac{\varepsilon_{t-k}^*}{\sqrt{h_t^*}} \right| \\ &\leq \frac{c}{\sqrt{T}} \sum_{t=1}^{[Tr]} \left( \sum_{k=t}^{\infty} \widehat{\beta}^{(k-1)/2} \right) = \frac{c}{\sqrt{T}} \sum_{t=1}^{[Tr]} O\left(\widehat{\beta}^{t/2}\right) \rightarrow 0 \text{ as } T \rightarrow \infty, \end{aligned}$$

where  $c$  is a constant. ■

### A.2 Proofs of Lemmas and Theorems

PROOF OF LEMMA 2

PART (a): For any  $\eta \in \mathbb{R}$ ,

$$\begin{aligned} F_T^{sym}(\eta) &= \frac{1}{2T} \sum_{i=1}^{2T} I(\eta_i \leq \eta) = \frac{1}{2T} \sum_{i=1}^T I(\eta_i \leq \eta) + \frac{1}{2T} \sum_{i=1}^T I(-\eta_i \leq \eta) \\ &= \frac{1}{2} [F_T(\eta) + (1 - F_T(-\eta))] \longrightarrow \frac{1}{2} [F(\eta) + (1 - F(-\eta))] = F(\eta) \text{ as } T \rightarrow \infty \end{aligned} \tag{6}$$

and by the symmetry of  $F$ . Because  $d_2$  is a metric,  $d_2 \left( \tilde{F}_T^{sym}, F \right)^2 \leq d_2 \left( \tilde{F}_T^{sym}, F_T^{sym} \right)^2 + d_2 \left( F_T^{sym}, F \right)^2$  and  $d_2 \left( F_T^{sym}, F \right)^2 \rightarrow 0$  from (6) and Bickel and Freedman (1981). Next, it is easy to show (Pascual *et al.*, 2000) that  $d_2(\tilde{F}_T^{sym}, F_T^{sym})^2 \leq E^* |\tilde{\eta}_j - \eta_j|^2 \leq \frac{6}{T} \sum_{j=1}^T (\tilde{\eta}_j - \eta_j)^2 + \frac{3}{T^2} \left( \sum_{i=1}^T \eta_i \right)^2 \rightarrow 0$  since  $T^{-1/2} \sum_{i=1}^T \eta_i = O_p(1)$  and  $\frac{6}{T} \sum_{j=1}^T (\tilde{\eta}_j - \eta_j)^2 = \frac{6}{T} \sum_{j=1}^T \frac{\varepsilon_j^2}{h_j \hat{h}_j} \left( \hat{h}_j^{1/2} - h_j^{1/2} \right)^2 \rightarrow 0$  from Auxiliary Lemma 1.

PART (b): Note that the  $k^{th}$  moment of the symmetrized residuals  $\{\tilde{\eta}_1, \tilde{\eta}_2, \dots, \tilde{\eta}_T, -\tilde{\eta}_1, -\tilde{\eta}_2, \dots, -\tilde{\eta}_T\}$  is given by  $(2T)^{-1} \left[ \sum_{t=1}^T (\tilde{\eta}_t)^k + \sum_{t=1}^T (-\tilde{\eta}_t)^k \right] = T^{-1} \sum_{t=1}^T (\tilde{\eta}_t)^p$  if  $p \geq 1$  is even and 0 if  $p$  is odd. Then, since  $\{\eta_t^*\}$  is an *iid* sample from the symmetrized empirical distribution function of the recentered standardized residuals,  $E^*(\eta_t^*) = 0$  and hence  $E^*(\varepsilon_t^*) = 0$ .

PART (c): Because  $E^*(\varepsilon_t^*) = 0$  and  $\eta_t^*$  are *iid* conditionally on the sample,  $Var^*(\varepsilon_t^*) = E^*(\varepsilon_t^*)^2 = E^*(h_t^*)E^*(\varepsilon_t^*)^2$  and  $Var^*(\varepsilon_t^*) = \left( \frac{1}{T} \sum_{t=1}^T \hat{h}_t \right) \left( \frac{1}{T} \sum_{t=1}^T \hat{\eta}_t^2 \right)$ . From part (a) of Auxiliary Lemma 1 and  $\sum_{t=1}^T \beta^t = O_p(1)$ , it follows that  $\frac{1}{T} \sum_{t=1}^T (\hat{h}_t - h_t) = o_p(1)$  and  $\frac{1}{T} \sum_{t=1}^T \hat{h}_t \xrightarrow{p} E(h_t)$ . Combining this result with  $\frac{1}{T} \sum_{t=1}^T \hat{\eta}_t^2 \xrightarrow{p} 1$ , we have  $Var^*(\varepsilon_t^*) \xrightarrow{p} \sigma^2$  as  $T \rightarrow \infty$ .

PART (d): Since  $E^*(\eta_t^*)^3 = 0$  by construction,  $\eta_t^*$  are *iid* conditionally on the sample, and  $\frac{1}{T} \sum_{t=1}^T \hat{h}_t = O_p(1)$ , we obtain that  $E^*(\varepsilon_t^*)^3 = 0$ . ■

### PROOF OF LEMMA 3

PART (a): By recursive substitution,

$$h_t^* = \hat{\omega} \left[ 1 + \sum_{k=1}^{t-1} \prod_{i=1}^k \left( \hat{\alpha} \eta_{t-i}^{*2} + \hat{\beta} \right) \right] + \hat{\omega} \prod_{i=1}^t \left( \hat{\alpha} \eta_{t-i}^{*2} + \hat{\beta} \right) h_0^*.$$

If the two candidate initial values are  $h_{01}^*$  and  $h_{02}^*$ , then the difference between the corresponding sequences  $h_{t1}^*$  and  $h_{t2}^*$  is given by  $|h_{t1}^* - h_{t2}^*| = \hat{\omega} \prod_{i=1}^t \left( \hat{\alpha} \eta_{t-i}^{*2} + \hat{\beta} \right) |h_{01}^* - h_{02}^*|$  and

$$E^* |h_{t1}^* - h_{t2}^*| = \hat{\omega} |h_{01}^* - h_{02}^*| E^* \left[ \prod_{i=1}^t \left( \hat{\alpha} \eta_{t-i}^{*2} + \hat{\beta} \right) \right] = \hat{\omega} |h_{01}^* - h_{02}^*| \left( \hat{\alpha} + \hat{\beta} \right)^t,$$

using that  $E^*(\eta_t^*)^2 \rightarrow 1$ . Since  $\hat{\alpha} + \hat{\beta} < 1$  by construction,  $E^* |h_{t1}^* - h_{t2}^*| \rightarrow 0$  as  $t \rightarrow \infty$ .

PART (b): Rewrite  $|\sqrt{h_{t1}^*} - \sqrt{h_{t2}^*}|$  as  $|\sqrt{h_{t1}^*} - \sqrt{h_{t2}^*}| = \left| \frac{h_{t1}^* - h_{t2}^*}{\sqrt{h_{t1}^*} + \sqrt{h_{t2}^*}} \right| \leq \left| \frac{h_{t1}^* - h_{t2}^*}{2\hat{\omega}} \right|$ . Then,

$$\begin{aligned} E^* |\varepsilon_{t1}^* - \varepsilon_{t2}^*| &= E \left[ \left| \sqrt{h_{t1}^*} - \sqrt{h_{t2}^*} \right| |\eta_t^*| \right] \\ &\leq E^* \left( \left| \frac{h_{t1}^* - h_{t2}^*}{2\hat{\omega}} \right| |\eta_t^*| \right) = \frac{1}{2} |h_{01}^* - h_{02}^*| \left( \hat{\alpha} + \hat{\beta} \right)^t E^* |\eta_t^*| \end{aligned}$$

using that  $E^*(\eta_t^*)^2 \rightarrow 1$ . Since  $E^*|\eta_t^*| < \infty$  and  $\hat{\alpha} + \hat{\beta} < 1$ ,  $E|\varepsilon_{t1}^* - \varepsilon_{t2}^*| \rightarrow 0$  as  $t \rightarrow \infty$ .

PART (c): Note that

$$\begin{aligned} E^* \left| \frac{1}{\sqrt{T}} \sum_{i=1}^T \varepsilon_{i1}^* - \frac{1}{\sqrt{T}} \sum_{i=1}^T \varepsilon_{i2}^* \right| &\leq \frac{1}{\sqrt{T}} \sum_{i=1}^T E^* |\varepsilon_{i1}^* - \varepsilon_{i2}^*| \leq \frac{1}{\sqrt{T}} \sum_{i=1}^T \frac{1}{2} |h_{01}^* - h_{02}^*| (\hat{\alpha} + \hat{\beta})^i E^* |\eta_t^*| \\ &= \frac{|h_{01}^* - h_{02}^*| E^* |\eta_t^*|}{2\sqrt{T}} \frac{1 - (\hat{\alpha} + \hat{\beta})^{T+1}}{1 - (\hat{\alpha} + \hat{\beta})} \leq \frac{1}{2\sqrt{T}} \frac{|h_{01}^* - h_{02}^*| E^* |\eta_t^*|}{1 - (\hat{\alpha} + \hat{\beta})} = O\left(\frac{1}{\sqrt{T}}\right). \end{aligned}$$

PART (d): Taking expectations under  $P^*$  of the difference between  $S_{[Tr]}^{(1)}$  and  $S_{[Tr]}^{(2)}$  yields

$$\begin{aligned} E^* \left| S_{[Tr]}^{(1)} - S_{[Tr]}^{(2)} \right| &\leq \frac{\lambda_1}{\sqrt{T}} E^* \sum_{t=1}^{[Tr]} |\varepsilon_{t1}^* - \varepsilon_{t2}^*| + \frac{\lambda_2}{\sqrt{T}} E \sum_{t=1}^{[Tr]} \left| \frac{\varepsilon_{t1}^*}{h_{t1}^*} - \frac{\varepsilon_{t2}^*}{h_{t2}^*} \right| \\ &\quad + \frac{\lambda_2 \hat{\alpha}}{\sqrt{T}} \sum_{t=1}^{[Tr]} E^* \left| \sum_{k=1}^{t-1} \hat{\beta}^{k-1} \frac{\varepsilon_{t-k1}^*}{h_{t1}^*} - \sum_{k=1}^{t-1} \hat{\beta}^{k-1} \frac{\varepsilon_{t-k2}^*}{h_{t2}^*} \right| E^* |(\eta_t^{*2} - 1)| \\ &= \lambda_1 I_1 + \lambda_2 I_2 + \lambda_2 \hat{\alpha} I_3. \end{aligned}$$

From part (c) we know  $I_1 = O(T^{-1/2})$ . Furthermore,  $I_2 = \frac{1}{\sqrt{T}} E^* \sum_{t=1}^{[Tr]} \left| \left( \frac{1}{\sqrt{h_{t1}^*}} - \frac{1}{\sqrt{h_{t2}^*}} \right) \eta_t^* \right| \leq \frac{1}{\hat{\omega}\sqrt{T}} E^* \sum_{t=1}^T |\varepsilon_{t1}^* - \varepsilon_{t2}^*| = O(T^{-1/2})$ . Using similar arguments, it can be shown (see Gospodinov and Tao, 2009) that  $I_3 = O(T^{-1/2})$  and, hence,  $E^* \left| S_{[Tr]}^{(1)} - S_{[Tr]}^{(2)} \right| = O(T^{-1/2})$ . ■

#### PROOF OF LEMMA 4

PART (a): As in part (a) of Auxiliary Lemma 1 and part (c) of Lemma 2, we can show that  $h_t^* - h_t = o_p(1) + O(\beta^t)$  and  $\frac{1}{T} \sum_{t=1}^T (h_t^* - h_t) = o_p(1)$  which implies that  $E^*(h_t^*) \rightarrow^p E(h_t)$ .

PART (b): Since both  $\varepsilon_t^*$  and  $h_t^*$  are stationary and ergodic (Auxiliary Lemma 2),  $\frac{1}{h_t^*}$  and  $\frac{\varepsilon_{t-k}^{*2}}{h_t^{*2}}$  are also stationary and ergodic. Using  $E^*(1/h_t^*) = \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t^*} + o_p(1)$  and  $E(1/h_t) = \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t} + o_p(1)$ , we have

$$\begin{aligned} \left| \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t^*} - \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t} \right| &= \left| \frac{1}{T} \sum_{t=1}^T \frac{\hat{h}_t - h_t^*}{h_t^* \hat{h}_t} \right| \leq \frac{1}{\omega \hat{\omega}} \left| \frac{1}{T} \sum_{t=1}^T (h_t - h_t^*) \right| \\ &= \frac{1}{\omega \hat{\omega}} |E h_t - E^*(h_t^*) + o_p(1)| = o_p(1). \end{aligned}$$

PART (c): From part (b) we already have that  $\left| \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t} - \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t^*} \right| = o_p(1)$ . Next,

$$\begin{aligned}
& \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^{*2}/h_t^{*2} - \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^2/h_t^2 \right| \\
&= \left| \frac{1}{T} \sum_{t=1}^T \frac{h_t^2 (\varepsilon_{t-k}^{*2} - \varepsilon_{t-k}^2) + (h_t^2 - h_t^{*2}) \varepsilon_{t-k}^2}{h_t^2 h_t^{*2}} \right| \\
&\leq \frac{1}{T \widehat{\omega}^2} \left| \sum_{t=1}^T (\varepsilon_{t-k}^{*2} - \varepsilon_{t-k}^2) \right| + \frac{1}{T \widehat{\omega}} \left| \sum_{t=1}^T \frac{(h_t + h_t^*)(h_t - h_t^*)}{h_t h_t^*} \frac{\varepsilon_{t-k}^2}{h_t} \right| \\
&\leq \frac{1}{\widehat{\omega}^2} \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^{*2} - \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^2 \right| + \frac{(\omega + \widehat{\omega}) \beta^{-(k-1)}}{\omega \widehat{\omega}^2} \left| \frac{1}{T} \sum_{t=1}^T h_t - \frac{1}{T} \sum_{t=1}^T h_t^* \right| \\
&= \frac{1}{\widehat{\omega}^2} |E^*(\varepsilon_{t-i}^{*2}) - E(\varepsilon_{t-i}^2) + o_p(1)| + \frac{(\omega + \widehat{\omega}) \beta^{-(k-1)}}{\omega \widehat{\omega}^2} |E(h_t) - E^*(h_t^*) + o_p(1)| \\
&= \beta^{-(k-1)} o_p(1) \text{ as } T \rightarrow \infty.
\end{aligned}$$

Then,

$$\begin{aligned}
|K - K^*| &\leq \left| \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t} - \frac{1}{T} \sum_{t=1}^T \frac{1}{h_t^*} \right| + \left| (\kappa \alpha^2 - \kappa^* \hat{\alpha}^2) \sum_{k=1}^{\infty} \beta^{2(k-1)} E(\varepsilon_{t-k}^2/h_t^2) \right| \\
&\quad + \left| \kappa^* \hat{\alpha}^2 \left( \sum_{k=1}^{\infty} \beta^{2(k-1)} \left( \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^2/h_t^2 + o_p(1) \right) - \sum_{k=1}^{\infty} \hat{\beta}^{2(k-1)} \left( \frac{1}{T} \sum_{t=1}^T \varepsilon_{t-k}^{*2}/h_t^{*2} \right) \right) \right| + o_p(1) \\
&= K_1 + |\kappa \alpha^2 - \kappa^* \hat{\alpha}^2| K_2 + \kappa^* \hat{\alpha}^2 K_3.
\end{aligned}$$

The first term  $K_1$  is  $o_p(1)$  (see part (b) above) and from the results of Lemma 2 and the properties of the MLE, it follows that  $|\kappa \alpha^2 - \kappa^* \hat{\alpha}^2| = o_p(1)$ . Furthermore,

$$\begin{aligned}
K_3 &= \left| \sum_{k=1}^{\infty} \beta^{2(k-1)} \left( \frac{1}{T} \sum_{t=1}^T (\varepsilon_{t-k}^2/h_t^2 - \varepsilon_{t-k}^{*2}/h_t^{*2}) \right) + \sum_{k=1}^{\infty} \left( \frac{1}{T} \sum_{t=1}^T (\beta^{2(k-1)} - \hat{\beta}^{2(k-1)}) \varepsilon_{t-k}^{*2}/h_t^{*2} \right) \right| \\
&\leq \frac{1}{\widehat{\omega}} \left| \sum_{k=1}^{\infty} \left( \left( \frac{\beta^2}{\hat{\beta}} - \hat{\beta} \right) \left( \left( \frac{\beta^2}{\hat{\beta}} \right)^{(k-2)} + \left( \frac{\beta^2}{\hat{\beta}} \right)^{(k-3)} \hat{\beta} + \dots + \hat{\beta}^{(k-2)} \right) \right) \right| + o_p(1) \\
&\leq \frac{1}{\widehat{\omega}} \left| \sum_{k=1}^{\infty} \left( \frac{\beta^2}{\hat{\beta}} - \hat{\beta} \right) (k-1) \lambda^{(k-2)} \right| + o_p(1) = o_p(1) \lim_{T \rightarrow \infty} \left( \frac{1 - \lambda^{T-1}}{1 - \lambda} + (T-1) \lambda^{T-1} \right) + o_p(1) = o_p(1),
\end{aligned}$$

where  $\lambda = \beta^2/\hat{\beta} < 1$  (for more details, see Gospodinov and Tao, 2009). Therefore,  $|K - K^*| = o_p(1)$  as  $T \rightarrow \infty$ .

PART (d): We can use similar arguments as in part (c) to prove  $F^* \xrightarrow{p} F$ . ■



# PROOF OF LEMMA 5

The structure of the proof is similar to that of Lemma 3 in Gospodinov (2008) for two-parameter partial sum processes and Lemma 4.2 in Ling and Li (2003). For full details, see Gospodinov and Tao (2009). ■

# PROOF OF THEOREM 1

Following Ling and Li (2003), the first two derivatives of the likelihood for observation  $t$  with respect to  $\phi$  can be expressed as

$$\begin{aligned} \left[ \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t^*(\phi, \delta^*)}{\partial \phi} \right]_{\phi=1} &= \frac{1}{T} \sum_{t=1}^T y_{t-1}^* \left[ \frac{\varepsilon_t^*}{h_t^*} + (1 - \eta_t^{*2}) \frac{\hat{\alpha}}{h_t^*} \sum_{j=1}^t \hat{\beta}^{j-1} \varepsilon_{t-j}^* \right] + o_p(1) \\ - \left[ \frac{1}{T^2} \sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta^*)}{\partial \phi^2} \right]_{\phi=1} &= \left[ T^{-2} \sum_{t=1}^T y_{t-1}^{*2} \right] \left[ E^* \left( \frac{1}{h_t^*} \right) + 2\hat{\alpha}^2 \frac{\varepsilon_t^{*2}}{h_t^*} \sum_{j=1}^t \hat{\beta}^{2(j-1)} E^* \left( \frac{\varepsilon_{t-j}^{*2}}{h_t^{*2}} \right) \right] + o_p(1). \end{aligned}$$

From Lemmas 4 and 5 and the continuous mapping theorem,  $\left[ \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t^*(\phi, \delta^*)}{\partial \phi} \right]_{\phi=1} \Rightarrow \int_0^1 W_1(r) dW_2(r)$  and  $-\left[ \frac{1}{T^2} \sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta^*)}{\partial \phi^2} \right]_{\phi=1} \Rightarrow F \int_0^1 W_1(r)^2 dr$ . Thus,

$$\left[ - \sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta^*)}{\partial \phi^2} \right]_{\phi=\phi_{ML}^*}^{1/2} (\phi_{ML}^* - 1) \Rightarrow \frac{1}{\sqrt{F}} \frac{\int_0^1 W_1(r) dW_2(r)}{\left( \int_0^1 W_1(r)^2 dr \right)^{1/2}}.$$

Define  $W_1(r) = \sigma B_1(r)$ ,  $\rho = 1/\sigma\sqrt{K}$  and  $W_2(r) = \sqrt{K}[\rho B_1(r) + \sqrt{1-\rho^2} B_2(r)]$ , where  $B_1(r)$  and  $B_2(r)$  are two independent standard Brownian motions. Substituting for  $W_1(r)$  and  $W_2(r)$  in the above expression, we get

$$\left[ - \sum_{t=1}^T \frac{\partial^2 l_t^*(\phi, \delta^*)}{\partial \phi^2} \right]_{\phi=\phi_{ML}^*}^{1/2} (\phi_{ML}^* - 1) \Rightarrow \sqrt{\frac{K}{F}} \left[ \rho \frac{\int_0^1 B_1(r) dB_2(r)}{\left( \int_0^1 B_1^2(r) dr \right)^{1/2}} + \sqrt{1-\rho^2} z \right] \quad (7)$$

noting that  $\left( \int_0^1 B_1^2(r) dr \right)^{-1/2} \int_0^1 B_1(r) dB_2(r)$  is distributed as a standard normal random variable  $z$ . From (7) and Polya's theorem, we obtain the desired result

$$\sup_{x \in \mathbb{R}} \left| P^*(t_{\phi_{ML}=1}^* \leq x) - P(t_{\phi_{ML}=1} \leq x) \right| \xrightarrow{P} 0.$$

■

Table 1. Empirical size (in %) of unit root tests (standard normal errors and  $T = 200$ ).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	1.00	5.05	9.35	1.31	5.80	9.95	1.35	5.40	10.06	1.05	5.35	9.85
$\alpha=0.5, \beta=0.4$	3.20	10.16	15.46	1.10	5.70	11.31	1.95	8.05	15.71	1.00	5.05	10.51
$\alpha=0.25, \beta=0.7$	2.85	8.70	14.21	0.85	5.15	10.11	2.15	7.75	15.21	1.10	5.10	9.50
$\alpha=0.399, \beta=0.6$	29.61	43.17	50.58	3.55	7.85	13.96	9.35	26.41	39.82	1.40	4.95	10.05
$\alpha=0.199, \beta=0.8$	13.51	25.66	35.07	1.95	6.70	11.71	9.15	29.16	42.72	1.00	4.95	9.95
$\alpha=0.7, \beta=0.25$	4.80	11.71	17.71	1.65	6.65	11.71	2.65	11.41	21.61	1.00	4.70	10.00
$\alpha=0.6, \beta=0.399$	31.83	45.95	53.55	4.20	11.16	16.66	8.91	24.02	37.04	1.50	5.51	10.11
$\alpha=0.8, \beta=0.199$	31.43	40.64	47.05	4.85	11.26	17.51	7.81	22.02	33.83	0.90	4.70	9.91

*Notes:* The empirical size is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim N(0, 1)$ ,  $\phi = 1$  and  $T = 200$ .

Table 2. Empirical size (in %) of unit root tests (standard normal errors and  $T = 400$ ).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	1.00	5.90	11.11	1.05	5.25	10.56	1.25	5.95	10.66	0.90	5.80	10.21
$\alpha=0.5, \beta=0.4$	2.75	8.05	13.26	1.50	5.25	10.46	1.55	6.85	13.16	1.10	4.95	9.45
$\alpha=0.25, \beta=0.7$	3.00	8.25	14.96	1.20	5.00	10.31	0.75	4.80	10.31	1.00	5.60	10.61
$\alpha=0.399, \beta=0.6$	28.86	40.22	47.02	2.15	6.80	13.36	3.35	14.46	25.01	1.40	5.65	10.96
$\alpha=0.199, \beta=0.8$	17.86	29.86	38.52	1.75	6.30	13.56	5.80	21.36	34.27	1.45	6.35	11.95
$\alpha=0.7, \beta=0.25$	4.75	11.31	16.86	0.80	4.50	9.70	2.10	8.65	17.36	1.00	4.75	9.30
$\alpha=0.6, \beta=0.399$	28.31	38.82	45.67	3.10	8.75	13.46	6.70	21.06	32.67	1.30	5.60	10.21
$\alpha=0.8, \beta=0.199$	25.56	34.17	39.72	3.65	9.45	15.51	6.25	18.81	30.12	1.45	4.75	9.05

*Notes:* The empirical size is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim N(0, 1)$ ,  $\phi = 1$  and  $T = 400$ .

Table 3. Empirical power (in %) of unit root tests (standard normal errors and  $T = 200$ ).

	<i>ASY - DF</i>			<i>BOOT - DF</i>			<i>ASY - GARCH</i>			<i>BOOT - GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	31.2	68.8	85.9	24.1	63.4	82.6	25.4	66.7	84.4	24.7	63.9	81.2
$\alpha=0.5, \beta=0.4$	15.9	48.7	67.7	23.8	55.9	72.2	56.9	87.5	94.7	52.8	86.3	93.5
$\alpha=0.25, \beta=0.7$	15.6	49.6	72.5	23.2	57.1	74.5	33.0	73.7	88.6	29.4	70.8	87.0
$\alpha=0.399, \beta=0.6$	0.2	2.4	9.6	4.2	15.6	27.2	22.2	63.7	76.2	28.7	61.3	73.4
$\alpha=0.199, \beta=0.8$	0.4	9.6	27.7	8.1	25.5	39.7	11.3	55.4	73.3	17.4	53.2	71.5
$\alpha=0.7, \beta=0.25$	6.5	42.2	63.4	22.1	51.5	69.5	74.5	94.5	97.3	68.4	91.7	96.5
$\alpha=0.6, \beta=0.399$	0.1	1.4	6.7	6.2	21.1	31.8	38.6	76.2	83.4	46.5	73.9	82.6
$\alpha=0.8, \beta=0.199$	0.0	2.2	7.1	10.9	28.8	40.6	70.6	89.0	92.6	63.4	84.6	90.3

*Notes:* The empirical power is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim N(0, 1)$ ,  $\phi = 0.92$  and  $T = 200$ . The power reported for the asymptotic tests (*ASY - DF* and *ASY - GARCH*) is size-adjusted power and the power for the bootstrap tests (*BOOT - DF* and *BOOT - GARCH*) is raw power.

Table 4. Empirical power (in %) of unit root tests (standard normal errors and  $T = 400$ ).

	<i>ASY - DF</i>			<i>BOOT - DF</i>			<i>ASY - GARCH</i>			<i>BOOT - GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	94.5	99.6	100	88.6	99.3	100	91.2	99.4	99.7	86.9	99.3	99.8
$\alpha=0.5, \beta=0.4$	67.6	94.8	98.3	71.6	92.5	96.4	99.4	100	100	98.5	99.9	100
$\alpha=0.25, \beta=0.7$	70.6	96.1	99.2	74.6	95.1	98.7	95.1	99.6	100	92.7	99.5	100
$\alpha=0.399, \beta=0.6$	0.4	12.9	39.5	16.9	44.9	60.7	83.5	94.1	96.4	81.5	93.6	96.2
$\alpha=0.199, \beta=0.8$	1.0	42.7	70.1	23.7	54.8	73.2	72.6	93.1	96.1	7.6	93.5	96.4
$\alpha=0.7, \beta=0.25$	47.5	87.8	95.7	65.5	87.8	94.5	99.6	99.9	100	99.2	99.8	100
$\alpha=0.6, \beta=0.399$	0.3	8.0	26.9	24.1	49.7	63.3	93.7	97.2	98.1	88.9	96.0	97.5
$\alpha=0.8, \beta=0.199$	0.2	6.6	22.7	35.1	61.1	72.5	96.8	98.3	98.8	90.4	97.0	98.3

*Notes:* The empirical power is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim N(0, 1)$ ,  $\phi = 0.92$  and  $T = 400$ . The power reported for the asymptotic tests (*ASY - DF* and *ASY - GARCH*) is size-adjusted power and the power for the bootstrap tests (*BOOT - DF* and *BOOT - GARCH*) is raw power.

Table 5. Empirical size (in %) of unit root tests (standardized  $t_7$  distribution).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	1.00	4.85	9.00	0.70	4.80	10.11	1.20	5.15	9.90	1.10	4.70	10.21
$\alpha=0.5, \beta=0.4$	3.85	9.20	15.11	1.05	5.90	11.26	3.50	11.01	18.06	1.30	5.90	10.81
$\alpha=0.25, \beta=0.7$	3.55	9.00	14.31	1.25	4.55	9.70	1.45	6.15	13.31	1.30	5.50	10.01
$\alpha=0.399, \beta=0.6$	31.87	44.67	53.28	3.35	7.90	13.01	7.00	21.16	35.51	1.70	6.70	12.06
$\alpha=0.199, \beta=0.8$	17.56	31.57	41.27	2.10	6.20	11.26	7.25	25.86	40.52	2.05	5.95	10.96
$\alpha=0.7, \beta=0.25$	5.45	12.41	17.71	1.50	6.70	11.81	5.00	14.51	23.46	1.45	5.45	10.76
$\alpha=0.6, \beta=0.399$	31.93	44.1	50.56	4.20	11.62	18.02	12.21	26.43	37.04	2.00	6.81	11.01
$\alpha=0.8, \beta=0.199$	30.83	39.14	44.24	6.11	12.91	19.12	10.41	25.23	36.94	1.80	6.70	10.01

*Notes:* The empirical size is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim t_7\sqrt{5/7}$ ,  $\phi = 1$  and  $T = 200$ .

Table 6. Empirical power (in %) of unit root tests (standardized  $t_7$  distribution).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	29.5	68.9	87.0	27.1	65.7	83.5	25.1	65.1	82.5	22.5	60.2	80.1
$\alpha=0.5, \beta=0.4$	11.3	47.5	70.2	23.0	56.0	72.6	36.9	73.4	88.6	40.4	76.8	88.5
$\alpha=0.25, \beta=0.7$	14.2	46.9	71.1	21.8	54.4	71.7	18.5	62.0	80.1	21.9	60.0	78.1
$\alpha=0.399, \beta=0.6$	0.5	1.4	6.9	3.3	14.4	25.9	12.3	46.7	63.8	21.3	51.5	66.5
$\alpha=0.199, \beta=0.8$	0.5	5.6	21.1	6.5	20.6	33.3	2.8	39.2	58.8	12.7	43.0	60.2
$\alpha=0.7, \beta=0.25$	6.9	35.7	61.1	21.9	51.5	67.6	49.7	83.8	93.4	54.5	84.1	92.8
$\alpha=0.6, \beta=0.399$	0.2	2.1	5.0	7.0	21.1	32.6	24.6	65.0	78.6	38.3	66.9	78.9
$\alpha=0.8, \beta=0.199$	0.2	2.3	9.0	10.9	28.5	42.1	40.4	79.1	89.9	52.8	78.0	88.9

*Notes:* The empirical power is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim t_7\sqrt{5/7}$ ,  $\phi = 0.92$  and  $T = 200$ . The power reported for the asymptotic tests (*ASY – DF* and *ASY – GARCH*) is size-adjusted power and the power for the bootstrap tests (*BOOT – DF* and *BOOT – GARCH*) is raw power.

Table 7. Empirical size (in %) of unit root tests (standardized  $\chi^2_5$  distribution).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	1.50	5.55	10.11	1.65	5.70	10.66	2.45	6.90	12.36	1.45	5.60	10.71
$\alpha=0.5, \beta=0.4$	3.80	10.21	15.11	1.55	6.05	12.01	5.25	13.21	21.41	1.75	6.00	11.06
$\alpha=0.25, \beta=0.7$	2.85	8.55	14.51	0.85	5.35	11.31	1.85	9.50	16.11	1.55	6.40	11.31
$\alpha=0.399, \beta=0.6$	32.97	47.67	55.03	4.20	10.56	16.46	9.60	25.16	36.77	2.25	6.95	13.21
$\alpha=0.199, \beta=0.8$	17.22	30.73	39.24	1.90	7.71	14.11	11.41	27.83	39.74	2.00	7.00	13.31
$\alpha=0.7, \beta=0.25$	5.80	12.46	18.21	2.00	6.35	12.56	6.95	17.71	26.91	1.45	6.15	10.71
$\alpha=0.6, \beta=0.399$	36.74	47.25	53.95	6.51	13.21	19.42	15.82	31.83	42.24	1.50	6.71	12.01
$\alpha=0.8, \beta=0.199$	33.63	43.14	48.55	7.31	15.22	21.42	16.12	31.13	40.44	1.80	7.81	12.71

*Notes:* The empirical size is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim (\chi^2_5 - 5)/\sqrt{10}$ ,  $\phi = 1$  and  $T = 200$ .

Table 8. Empirical power (in %) of unit root tests (standardized  $\chi^2_5$  distribution).

	<i>ASY – DF</i>			<i>BOOT – DF</i>			<i>ASY – GARCH</i>			<i>BOOT – GARCH</i>		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\alpha=0, \beta=0$	19.0	62.9	85.5	27.6	66.1	83.8	12.3	58.2	77.9	20.4	60.3	79.2
$\alpha=0.5, \beta=0.4$	14.8	53.6	73.4	27.0	60.5	75.9	26.1	63.7	81.6	30.2	66.2	82.1
$\alpha=0.25, \beta=0.7$	15.7	57.9	78.1	24.1	61.5	78.4	12.1	45.9	66.1	16.8	49.2	69.3
$\alpha=0.399, \beta=0.6$	0.0	2.2	10.0	4.9	12.4	25.7	9.7	40.8	60.8	19.5	50.3	68.1
$\alpha=0.199, \beta=0.8$	0.2	8.7	24.4	5.1	21.3	35.2	4.10	31.2	51.1	11.8	40.0	59.7
$\alpha=0.7, \beta=0.25$	5.8	41.1	67.3	26.4	57.7	73.4	42.4	76.3	89.2	44.3	76.1	88.3
$\alpha=0.6, \beta=0.399$	2.0	2.9	8.1	6.4	20.2	32.3	26.6	61.0	75.0	33.1	63.2	75.1
$\alpha=0.8, \beta=0.199$	0.0	1.5	7.8	13.6	29.4	44.1	52.0	72.0	82.4	46.9	75.0	84.2

*Notes:* The empirical power is computed from 2,000 Monte Carlo replications with data generated from model (5) with  $\eta_t \sim (\chi^2_5 - 5)/\sqrt{10}$ ,  $\phi = 0.92$  and  $T = 200$ . The power reported for the asymptotic tests (*ASY – DF* and *ASY – GARCH*) is size-adjusted power and the power for the bootstrap tests (*BOOT – DF* and *BOOT – GARCH*) is raw power.

Table 9. Unit root tests for U.S. interest rates.

	<i>ASY – DF</i>		<i>BOOT – GARCH</i>	
	<i>test</i>	<i>p-value</i>	<i>test</i>	<i>p-value</i>
Fed funds rate	-2.052	0.265	-5.382	0.000
3-month rate	-1.994	0.290	-3.582	0.008
1-year rate	-1.957	0.306	-3.020	0.023
5-year rate	-1.629	0.468	-2.979	0.032
10-year rate	-1.578	0.494	-2.662	0.068
default premium	-1.329	0.618	-4.650	0.014

*Notes:* The  $p$ -values for the  $ASY – DF$  test are computed as in MacKinnon (1996). The  $p$ -values for the  $BOOT – GARCH$  test are obtained using the bootstrap procedure described in Section 3.1 with 1,999 bootstrap replications.

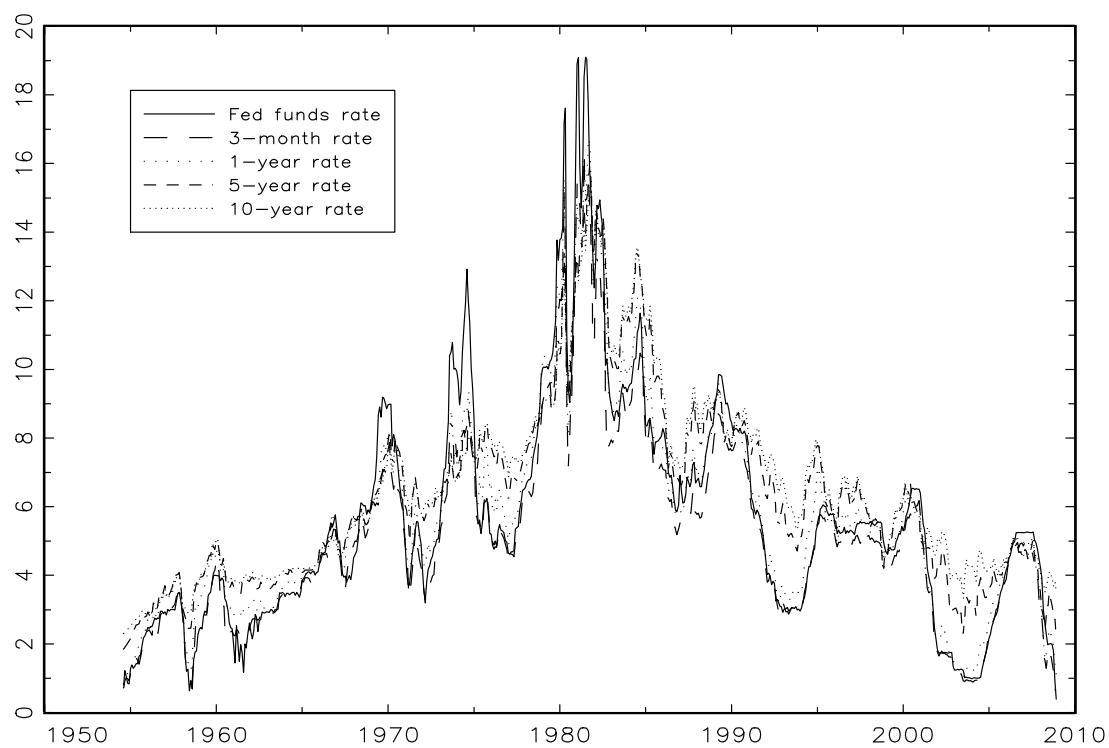


Figure 1. U.S. interest rates.

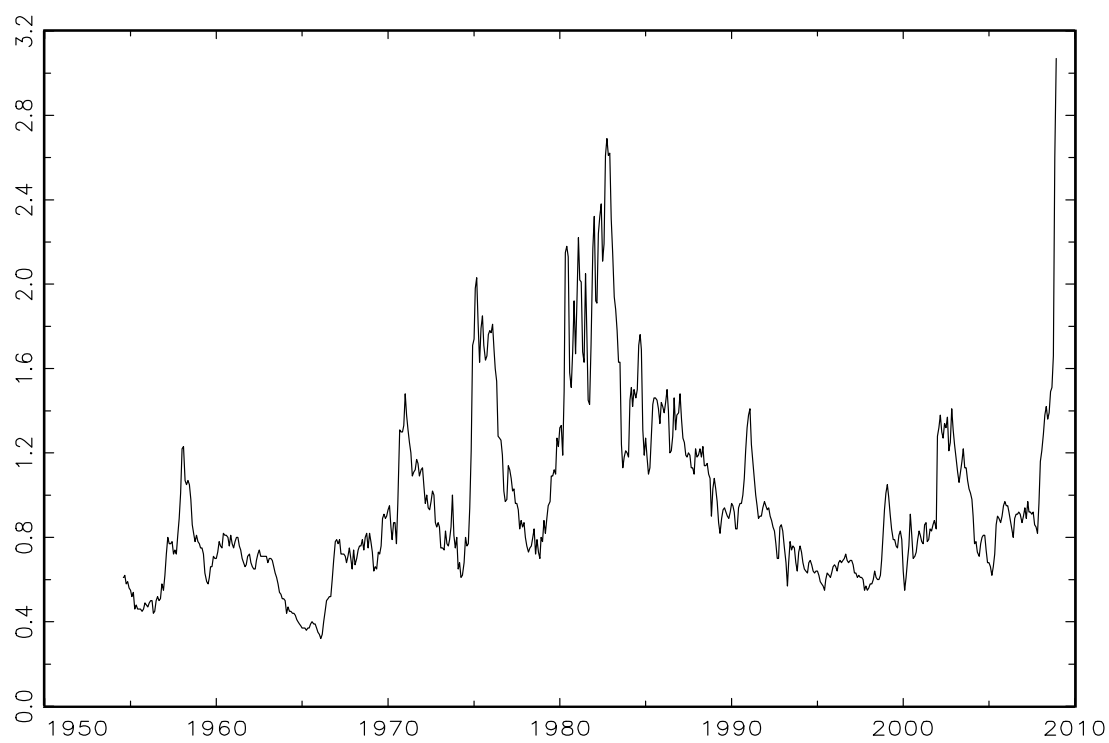


Figure 2. U.S. default premium (Baa corporate bond yield - Aaa corporate bond yield).